ORIGINAL RESEARCH PAPER
Volume - 12 | Issue - 07 | July - 2023 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

# THE PERCENT FRAGILITY INDEX

Biomedical

**Thomas F Heston** MD, FAAFP, FASNC, FACNM

## ABSTRACT

This article proposes the Percent Fragility Index (PFI) as an improved measure of statistical fragility in biomedical research. The PFI quantifies the percentage change in outcomes needed to change a study's statistical significance from positive to negative or vice-versa. The PFI improves upon existing indices by providing an intuitive statistic that is easy to grasp and by accommodating both dichotomous and continuous variables. This approach minimizes dependency. on sample size, a limitation of the commonly used Fragility Index (FI) and Fragility Quotient (FQ). The FI measures the minimum number of outcome events required to reverse statistical significance, and the FQ divides the FI by the total sample size. The PFI enhances the interpretability and validity of fragility assessments. PFI facilitates a more critical understanding of research outcomes by offering readers a more precise estimate of study fragility.

## KEYWORDS

## INTRODUCTION

There are currently several ways of quantifying statistical fragility in biomedical research. The first is the Unit Fragility Index (uFI), which quantifies the effect of small changes in clinical outcomes on the p-value (1,2). The uFI was then modified slightly to become the more commonly used Fragility Index (FI), an integer value representing the absolute number of outcome events required to reverse the statistical significance findings (3). Both metrics utilize the same concept of quantifying the effect of small changes in outcomes upon significance testing.

Another common statistic quantifying fragility is the fragility quotient (FQ), which is the FI divided by the total sample size (4). This statistic attempts to overcome the dependency of the FI upon sample size. A small change has a much greater effect on p-values than a similar change in a large sample size. By dividing the FI by the sample size, this dependency is minimized.

Because the FI does not vary based on sample size, it can be challenging to interpret in isolation. One way to put the FI in context is to compare it with the number of subjects lost to follow-up. The study is considered highly fragile when the FI is less than the number lost to follow-up. Another way to put the FI in context is to compare it with the number of unanalyzed patients. Again, if the FI is lower than the number of patients not analyzed, there is a high risk of losing significance if the study were repeated (5,6).

Another extension of the FI is the continuous FI (CFI). Whereas the FI is only used for dichotomous outcomes, the CFI is used for continuous outcomes. One unit from the higher mean is moved to the lower mean until the p-value exceeds 0.05. For example, if researchers are looking at the effect of a medication on cholesterol levels, and the mean cholesterol level in the intervention group is higher than the placebo group, the CFI is the number of unit changes from the higher mean to the lower mean it takes to make the comparison statistically insignificant (7)

Here I propose the Percent Fragility Index (FPI) to provide an improved quantitative measure of a study's fragility. The FPI is conceptually easy to understand and takes into account the sample size. Thus, integration into routine statistical analyses would provide readers with a quick and accurate assessment of the study's fragility.

### The Fragility Index

The FI looks specifically at the effect of iteratively changing outcomes of a binomial variable. This can be demonstrated by the use of a 2 x 2 contingency table. The FI is always an integer representing the minimum number of outcomes that would reverse the statistical significance of a particular outcome if changed. A lower FI indicates greater fragility. There is no current consensus on a cut-off value for determining whether or not a study is fragile (3). For our purposes, in a 2 x 2 table with statistically significant findings, the FI is calculated by iteratively decreasing the largest cell value by 1 and adjusting all other cells accordingly to keep the marginal totals unchanged. This is done iteratively until the findings change from significant to insignificant.

For a 2 x 2 table where the findings are statistically insignificant, the FI is calculated by iteratively increasing the largest cell value by 1 and adjusting the other cells accordingly to keep the marginal totals unchanged until the findings become significant. A large FI supports the assertion that the findings are robust, and a small FI suggests the findings are fragile.

### The Fragility Quotient

The FQ is a simple calculation of the FI divided by the total sample size. It ranges from o to 1. There is no consensus on a cut-off value for the FQ to indicate whether a study is robust or fragile. A large FQ supports the assertion that the findings are robust, and a small FQ suggests the findings are fragile.

### The Percent Fragility Index

The percent fragility index (PFI) looks at percent changes rather than unit changes in the cells. For a statistically significant 2 x 2 contingency table, the PFI is calculated by incrementally decreasing the value of the cell with the largest value and correspondingly adjusting all other cells to keep the marginal totals fixed. For a 2 x 2 table that is statistically insignificant, the PFI is calculated by incrementally decreasing the value of the cell with the largest value and correspondingly adjusting all other cells to keep the marginal totals fixed. This process is continued until the statistical significance is changed from significant to insignificant or vice versa. The PFI does not rely on integer changes in outcomes so it can be applied to dichotomous or continuous variables. It is more resistant to changes in sample size than the FQ, which relies on the underlying FI. In addition, the FPI gives readers an intuitive grasp of how fragile the data is by providing the percent change in outcomes required to flip the significance. Like the FI, the PFI is calculated by iteratively changing the largest value, which is increased if the 2 x 2 table is statistically insignificant or decreased if the 2 x 2 table is statistically significant.

### Examples

Tables 1 to 3 show the FI, FQ, and PFI calculation methods. Tables 4 to 9 demonstrate the fragility of statistically significant 2 x 2 contingency tables. Tables 10 to 12 demonstrate the fragility of a 2 x 2 table with insignificant results. Chi-square testing was utilized for all significance tests.

### Clinical Impact

Statistics can be misleading, and the desire to find something significant can be overpowering to researchers hoping to discover a new treatment or a new diagnostic tool. Indeed, the strong preponderance of publishing significant as opposed to insignificant findings has been repeatedly demonstrated by studies from a broad group of researchers from various institutions (8,9). After all, who wants to discover something insignificant?

It's well past time for medical researchers to wean themselves off an over-reliance on the p-value by including in their statistical analyses a quantitative measure of fragility. Use of the PFI would provide a meaningful indication of a study's fragility. The PFI can help advance medical science by providing clinicians with an improved estimation

of the validity of research findings.

**Table 1.** The standard 2 x 2 contingency table has 4 outcomes placed into 4 cells labeled a, b, c, and d.

| | Disease + | Disease - | |
|---|---|---|---|
| Exposure + | a | b | a + b |
| Exposure - | c | d | c + d |
| | a + c | b + d | a + b + c + d |

**Table 2.** The FI is the absolute value of how many single unit changes it takes to convert a non-significant finding to a significant one or vice versa. Only cells a, b, c, and d are changed, and the marginal totals remain fixed. For example, iIf a is increased by 1, then b and c are decreased by 1, and d is increased by 1. If a decreases by 1, then b and c increase by 1, and d decreases by 1. The FI is always applied so that the largest cell value is incrementally decreased or increased by 1 until the significance changes from significant to insignificant or vice-versa.

| | Disease + | Disease - | |
|---|---|---|---|
| Exposure + | a +/- FI | b +/- FI | a + b |
| Exposure - | c +/- FI | d +/- FI | c + d |
| | a + c | b + d | a + b + c + d |

**Table 3.** The PFI looks at what happens when the cells are changed by a percentage instead of an integer value. The marginal totals remain fixed. For example, if a is increased by 5% [a +( a*0.05)], then b and c are decreased by (a*0.05), and d is increased by (a*0.05). If a is decreased by 5% [a - (a*0.05)], then b and c are increased by (a*0.05), and d is decreased by (a*0.05). Note that the PFI always is applied to the cell with the highest value. If, for example, cell c was the highest value, then the change in each cell would be +/- (c * PFI).

| | Disease + | Disease - | |
|---|---|---|---|
| Exposure + | a +/- (a * PFI) | b – (a *PFI) | a + b |
| Exposure - | c +/- (a * PFI) | d + (a * PFI) | c + d |
| | a + c | b + d | a + b + c + d |

**Table 4.** These findings are statistically significant (p = 0.048)

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 20 | 14 |
| Exposure - | 7 | 15 |

**Table 5.** These findings are insignificant (p = 0.15). The FI = 1. The FQ = 0.018 (1/56). The statistical findings observed in Table 4 are highly fragile, and the results are viewed with high skepticism.

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 19 | 15 |
| Exposure - | 8 | 14 |

**Table 6.** These findings are insignificant (p = 0.0501). The FPI = 0.15%. The FPI helps clarify just how fragile the findings are in Table 4 by showing that only a 0.15% change in a is required to change the findings from significant to insignificant.

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 19.97 | 14.03 |
| Exposure - | 7.03 | 14.97 |

**Table 7.** These findings are statistically significant (p = 0.00031)

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 25 | 12 |
| Exposure - | 7 | 23 |

**Table 8.** These findings are insignificant (p=0.10). The FI = 4. The FQ = 0.0598 (4/67). These findings may or may not be considered fragile. If the number of subjects lost to follow-up is > 4, then most would consider the findings fragile.

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 20 | 21 |
| Exposure - | 12 | 27 |

**Table 9.** These findings are insignificant. The PFI = 17%. Compare this to the FQ. Whereas the FQ value of 0.0598 lacks any intuitive meaning, the FPI shows that a change of 17% in outcomes is required to flip the findings from significant to insignificant.

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 20.6 | 20.4 |
| Exposure - | 11.4 | 27.6 |

**Table 10.** These findings are statistically insignificant (p = 0.234)

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 25 | 18 |
| Exposure - | 19 | 23 |

**Table 11.** These findings are significant (p=0.0224). The FI = 8. The FQ = 0.0941 (8/85). The observations in Table 6 are less fragile than the findings in Table 4.

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 17 | 26 |
| Exposure - | 27 | 15 |

**Table 12.** These findings are significant. The PFI = 31%, consistent with the significant findings in Table 10 being highly robust.

| | Disease + | Disease - |
|---|---|---|
| Exposure + | 20.6 | 20.4 |
| Exposure - | 11.4 | 27.6 |

## REFERENCES


1. Feinstein AR. The unit fragility index: an additional appraisal of "statistical significance" for a contrast of two proportions. J Clin Epidemiol. 1990;43(2):201–9. PMID: 2303850.
2. Walter SD. Statistical significance and fragility criteria for assessing a difference of two proportions. J Clin Epidemiol. 1991;44(12):1373–8. DOI: 10.1016/0895-4356(91)90098-t. PMID: 1753268.
3. Walsh M, Srinathan SK, McAuley DF, Mrkobrada M, Levine O, Ribic C, et al. The statistical significance of randomized controlled trial results is frequently fragile: a case for a Fragility Index. J Clin Epidemiol. 2014 Jun;67(6):622–8. DOI: 10.1016/j.jclinepi.2013.10.019. PMID: 24508144.
4. Ahmed W, Fowler RA, McCredie VA. Does sample size matter when interpreting the fragility index? Crit Care Med. 2016 Nov;44(11):e1142–3. DOI: 10.1097/CCM.0000000000001976.
5. Matics TJ, Khan N, Jani P, Kane JM. The fragility index in a cohort of pediatric randomized controlled trials. J Clin Med. 2017 Aug 14;6(8). DOI: 10.3390/jcm6080079. PMID: 28805717. PMCID: PMC5575581.
6. Gaudino M, Hameed I, Biondi-Zoccai G, Tam DY, Gerry S, Rahouma M, et al. Systematic evaluation of the robustness of the evidence supporting current guidelines on myocardial revascularization using the fragility index. Circ Cardiovasc Qual Outcomes. 2019 Dec 11;12(12):e006017. DOI: 10.1161/CIRCOUTCOMES.119.006017. PMID: 31822120.
7. Ho AK. The fragility index for assessing the robustness of the statistically significant results of experimental clinical studies. J Gen Intern Med. 2022 Jan;37(1):206–11. DOI: 10.1007/s11606-021-06999-9. PMID: 34357573. PMCID: PMC8739402.
8. Hopewell S, Loudon K, Clarke MJ, Oxman AD, Dickersin K. Publication bias in clinical trials due to statistical significance or direction of trial results. Cochrane Database Syst Rev. 2009 Jan 21;2009(1):MR000006. DOI: 10.1002/14651858.MR000006.pub3. PMID: 19160345. PMCID: PMC8276556.
9. Flint J, Cuijpers P, Horder J, Koole SL, Munafò MR. Is there an excess of significant findings in published studies of psychotherapy for depression? Psychol Med. 2015 Jan;45(2):439–46. DOI: 10.1017/S0033291714001421. PMID: 25062429. PMCID: PMC4301215.